

Tugas Pekan ke-11 *Machine Translation*

Batas pengumpulan: Jumat 4 Desember 2020, pukul 10.59 pagi, melalui LMS

Deskripsi

Lakukan implementasi pelatihan *alignment* dari teks paralel korpus yang Anda definisikan sendiri, terdiri dari **20** pasang kalimat. Teks paralel korpus tersebut berisi teks dengan **bahasa sumber adalah Bahasa Indonesia**, dan **bahasa target adalah salah satu bahasa daerah yang Anda kuasai**.

Program dan Analisis yang Harus Dibuat

1. Buatlah program untuk pelatihan alignment dari korpus paralel, menggunakan NLTK, dengan algoritma IBMModel1.
2. Lakukan pengukuran skor probability alignment 3 pasang kata.
3. Lakukan analisis terhadap hasil pengukuran skor probability *alignment*:
 - a. Perbandingan skor probability, mana yang lebih tinggi, mana yang lebih rendah, mengapa.
 - b. Penjelasan singkat kaitannya dengan metode IBMModel1.

File yang harus dikumpulkan:

1. Program, berisi pendefinisian algoritma dan korpus paralel (boleh dituliskan langsung di program/*hardcode*).
2. Laporan: 1 file pdf, maksimum panjang laporan adalah 1 halaman. Laporan berisi jawaban pertanyaan nomor 2 dan 3 yang menjelaskan analisis skor *probability alignment* yang diperoleh.

Detail penilaian:

- a. Program:
 - Kebenaran implementasi pelatihan alignment dengan NLTK metode IBMModel 1 serta pengukuran skor probability alignment [50 poin]
- b. Laporan:
 - kelengkapan jawaban nomor 2 dan 3 [10 poin]
 - analisis jawaban nomor 3.a [20 poin]
 - analisis jawaban nomor 3.b [20 poin]

Jika ada pertanyaan, silakan disampaikan melalui *channel* pekan_11_machine_translation di slack.